DOCUMENT RESUME

ED 473 410 EA 032 098

AUTHOR	Loveless, Tom
TITLE	Charter School Achievement and Accountability.
REPORT NO	PEPG/02-09
PUB DATE	2002-00-00
NOTE	43p.; Paper prepared for the Program on Education Policy and Governance, John F. Kennedy School of Government, Harvard University Conference, "Taking Account of Accountability: Assessing Politics and Policies" (Boston, MA, June 10-11, 2002).
AVAILABLE FROM	For full text: http://www.ksg.harvard.edu/pepg/pdf/TAConfPDF/PEPG02-09%20TA%20Loveless.pdf.
PUB TYPE	Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE	EDRS Price MF01/PC02 Plus Postage.
DESCRIPTORS	*Academic Achievement; *Accountability; *Charter Schools; Elementary Secondary Education; Instructional Effectiveness; Nontraditional Education; Outcomes of Education; Public Schools; School Choice; *School Effectiveness

ABSTRACT

This report presents findings of an analytical study on student achievement in charter schools and the effect of state accountability regimes on charter-school performance. Data for the study were gathered from a literature review and test scores from 1999 to 2001 of both charter schools and regular public schools. The data were collected from state departments of education and from websites maintained by state assessment programs. A total of 638 charter schools from 10 states were selected for the study. Demographic data were obtained from the Common Core of Data. The following are some of the findings of the study: (1) achievement is significantly lower in charter schools than in regular public schools; (2) charters serving at-risk students achieve at significantly lower levels than open-admissions charters; (3) charter achievement is stronger in reading than in math and stronger in 8th grade than in 4th and 10th grades; (4) no evidence exists to show that accountability policies produce higher achievement in charter schools; and (5) the policies of closing charters, imposing school sanctions, and requiring students to pass exit exams negatively affect charter schools' test scores. (Contains 14 tables and 37 endnotes, most of which contain references.) (WFA)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 473 410

Charter School Achievement and Accountability

Tom Loveless
Brown Center on Education Policy
The Brookings Institution

PEPG/02-09

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY
T. LOVELESS
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
1

DRAFT: DO NOT QUOTE OR CITE WITHOUT AUTHOR'S PERMISSION

Paper prepared for *Taking Account of Accountability: Assessing Politics and Policies*, Program on Education Policy and Governance, John F. Kennedy School of Government, Harvard University, June 10-11, 2002. The author thanks Paul DiPerna and John Coughlan for research assistance and Christopher Avery, Thomas Kane, and Clifton Winston for answering technical questions that arose during the study. All errors are the responsibility of the author.

EA032098

BEST COPY AVAILABLE

Charter schools rose to prominence in the 1990s with the backing of two powerful reform movements in education, the accountability and choice movements. Supporters argue that charter schools are more accountable than traditional public schools, that if charter schools do not perform, they are shut down. One mechanism of accountability involves parental choice. Charter schools' funding is tied to enrollment. A charter school must attract students and please parents, or its main source of revenue disappears. Another mechanism arises from charter schools' renewable license to operate (or "charter"), which is granted by an authorizing agency--a local school board, university, state department of education, or some other institution. Authorizers act as the state's agents, monitoring charter schools and determining whether the outcomes promised in the schools' charters are met. If a charter school fails to satisfy its authorizing agency, its charter can be revoked.

From 1991 to 2001, while state legislatures were passing charter legislation and charter schools proliferated across the country--from the first few charters in Minnesota in 1992 to approximately 2,400 in 2001—state legislatures were also establishing programs to hold all schools, traditional and charter, accountable for student learning. Most states now regularly test students on academic subjects and either reward or sanction schools based on the results. In 2002, President Bush signed the No Child Left Behind Act, which mandates annual testing in grades three through eight and state intervention when schools persistently fail. In addition, several states have passed laws holding students more accountable for learning, and some require students to pass an examination before graduating from high school. Students in the graduating class of 2008 will face exit exams in twenty-eight states.[1]

Charter schools exist in thirty-four states that have state testing programs.[2] Because charter schools and their students are not exempt from testing, state accountability systems may magnify or otherwise influence the accountability mechanisms inherent to the charter school concept. Even in schools using multiple measures of student learning, state tests become the authoritative indicator of student achievement.[3] Parents who find out that their charter school

has been labeled "failing" by state authorities may consider enrolling their children elsewhere. And such a school would undoubtedly face heightened scrutiny from its authorizing agency. Thus, it is reasonable to expect accountability systems to ratchet up the importance of test scores in charter schools.

This paper analyzes student achievement in charter schools and examines the effect of different state accountability regimes on charter school performance. The paper consists of five sections. I first review the literature on student achievement in charter schools and the effectiveness of state accountability systems. The second section describes the study's methods and data, with attention paid to problems encountered with the data and the strategies chosen to address them. The third section presents an analysis of academic achievement. I analyze charter school achievement in ten states using test scores from 1999 to 2001. The objective is to gauge how charter schools performed during this period compared to the average public school in the ten states. The fourth section investigates whether the charters' performance varied by accountability regime, recognizing, with only ten states' systems to compare—and the limitations of the data at hand—that the analysis is far from finely-grained. The fifth section discusses the findings and implications for future charter school research.

Research on Charter Schools, State Accountability Systems, and Student Achievement

How do students in charter schools perform on tests of academic achievement? The charter school experiment is barely a decade old. Research has been slow to accumulate because most charters, even today, are relatively new, they lack standardized test score data on which to compare achievement, and the test data from very small charters, which make up a large portion of the total number of charter schools operating, are sketchy at best.[4]

In a recent review of the research on charter school achievement, Miron and Nelson identified 15 studies meeting basic quality standards.[5] The studies took place in seven states and the District of Columbia and were conducted by state departments of education, as well as by independent researchers. Miron and Nelson assigned weights to each study according to its

methodological strength and the magnitude of effects. They summarize the overall results as "mixed or very slightly positive" for charter schools. Positive charter effects were found in some states and negative effects in others. The strongest positive findings were from Arizona, and the strongest negative findings from Michigan. The negative findings in Michigan contrast with Hoxby's 2001 study showing that MI charter schools are producing an effect of "all boats rising." Test score gains in both charters and regular public schools were significant in districts where charter schools were drawing at least 6% of enrollment. Hoxby attributed the achievement gains to charters' competitive effects.[6]

The mixed results for charter schools was seconded by another summary of the research, this one completed by Brian Gill and a team of researchers at RAND.[7] None of these researchers were able to identify policies that states should implement to enhance the academic effectiveness of charters, nor the policies that should be avoided. On the policy front, Miron and Nelson noted that neither the "strength" of charter laws--the regulatory freedom granted to charters—nor the stringency of charter oversight--the percentage of charter schools that have been closed by authorizers—seemed to explain why successful charters are more likely to be located in some states than in others. Charter school research is in its infancy, so more helpful policy guidance could emerge in the future.

Sparse research also exists on the effect of state accountability systems on student achievement. Several studies have investigated how educators respond to accountability efforts, exploring such questions as whether an emphasis on testing narrows the curriculum, increases the amount of classroom time devoted to test preparation, or alters school organization and decision making.[8] Others have examined the effect of exit exams on drop out rates.[9] Although it may seem intuitive that holding educators accountable for student achievement will improve the likelihood of its attainment, there is a dearth of empirical evidence that accountability systems actually raise test scores. This is somewhat surprising because a few states have been testing students and publicly releasing test scores for two to three decades. Historically, however, schools and districts utilized the data for their own purposes. Scores were not used for

accountability. Only recently have states taken command of assessment and started attaching consequences for schools that fail to improve test scores or fail to score at a targeted level. Like charters, accountability systems are simply too new and vary too much across states for a reliable body of research to yet emerge.

The most interesting evidence on accountability has come from a series of studies by John Bishop and researchers at Cornell University. Bishop has shown that countries with curriculum based external exit exams (CBEEEs) score higher on international math and science assessments.[10] The same pattern holds true for Canadian provinces; those employing CBEEEs exhibit higher test scores than the provinces that do not.[11] In the U.S., Bishop found positive achievement effects for the New York Regents program and the Michigan Merit Award Program. Michigan's program offers one-year $2500 scholarships to students who meet or exceed standards in reading, math, science, and writing. The tests include demanding material, and students who fall short do not face negative consequences. Thus, the program is not high stakes or predicated on students demonstrating minimum competency in basic subjects, two aspects of accountability systems that draw fire from critics.[12]

Bishop has also examined mixed systems, targeting both students and educators. He analyzed the 8th grade NAEP scores of states with different accountability regimes—for students, meeting basic course requirements, passing minimum competency exams, and passing CBEEEs; and for schools, receiving rewards or sanctions based on test scores. Students in CBEEE states (New York and North Carolina) exhibited the highest levels of achievement, with an advantage of .45 grade levels in math and science, followed by states that rewarded or sanctioned schools, with gains of .20 grade levels. Minimum competency tests had a positive but insignificant effect. Requiring particular courses for high school graduation had no effect.[13]

Bishop concluded that systems that combine student and school accountability hold the most promise for raising academic achievement, especially when performance on end-of-course exams or other curriculum-based tests are the outcomes that states measure and reward. Grissmer et al. arrived at a similar conclusion after analyzing NAEP gains registered in the

1990s. They concluded that the success of states such as Texas and North Carolina is at least partially attributable to well developed accountability policies.[14]

Despite the paucity of research on both charter schools and state accountability systems, then, there have been documented successes with the two reforms. Thus far, to the best of my knowledge, no study has investigated how charter schools interact with state accountability systems. Do charters and state accountability work together to boost learning? Are charter schools in states with several accountability policies doing better than charters in states with few or no accountability policies?[15]

There are several reasons to expect accountability policies to have a positive effect on charter schools. Publicly released test scores--along with well-publicized ratings, rewards, and sanctions--place charter schools and regular public schools on an equal playing field. Accountability systems clarify the outcomes on which charters and regular public schools will compete, thereby contributing to charter schools' transparency.[16] If charter school supporters are correct and charters enjoy fewer bureaucratic constraints than traditional public schools, charters should be able to respond more quickly than public schools to community demand for achievement—and to advertise a record of academic success more effectively to potential clients. Charter school authorizers who believe state ranking systems are important may also pressure charters to perform at high levels on state tests.

Accountability and charter policies could also work at cross purposes. An essential element of the charter school idea is that institutions should be free to fashion their own goals and to govern themselves independently in pursuit of outcomes that parents and educators mutually value. Accountability systems intrude on that process. By defining the types of learning that schools will produce—and by establishing a system that measures progress toward those attainments—accountability reduces the autonomy of charter schools. Schools with low test scores may be shut down, whether parents like them or not. Students who fail exit exams may be denied a diploma, whether or not parents are satisfied with the education their children have received.

Data and Methods

The study seeks to assess academic achievement in charter schools and to evaluate how well charters are performing under various forms of accountability. The study's basic analytical approach is to compare charter schools' test scores from 1999-2001 to those of regular public schools, employing controls for student background characteristics that may influence achievement. After adjusting test scores for these factors, I examine whether charter school achievement varies by state accountability policies.

Ten states were selected for the study. The states had at least 30 charter schools open in 1999 and tested students in grades, 4, 8, and 10 (allowing for substitution of adjacent grade levels) using the same achievement test in 1999, 2000, and 2001. The ten states yield a total of 638 charter schools that were open in 1999. I did not add charters opening in 2000 or 2001 to the sample so that the number of schools remains fixed during the three year time frame. Test data were collected--on the charters and regular public schools--from state departments of education and from websites maintained by state assessment programs. For each school, a single composite score was computed combining reading and math scores over the three years. Within each state, a z-score was calculated for every school, standardizing school achievement with a mean of zero and standard deviation of one. Demographic data were obtained from the Common Core of Data, a database collected by the National Center for Education Statistics.

[TABLE 1 About Here]

Table 1 lists several technical challenges encountered while conducting the study, the strategies chosen to address them, and the limitations of the choices. I am not a methodologist. The study is not intended to blaze any new methodological trails nor are the solutions to technical issues very sophisticated. To the contrary, the principal goal is to simplify a complex body of evidence so that it will be accessible to interested observers with a basic understanding of statistics. I review the technical issues here to inform readers of important decisions that

were made at several stages of the analysis. These decisions reflect certain assumptions, affect the study's statistical results, and could influence the paper's conclusions.

States do not employ standard measures of achievement. Even in the same subject, states administer tests with vastly different content, assessment protocols, and methods of reporting scores. Computing z-scores for school in effect sets up a "horse race" within each state. Each school's achievement is expressed relative to its state's average score. Test scores that are above average convert to positive z-scores, those below average to negative z-scores. The benefit of this approach is that it expresses the achievement of schools from different states in the same metric. This works well for comparing schools within any particular group, but, being relative rankings, z-scores do not indicate the *direction* of the group as a whole. Scores can go up or down, and as long as schools maintain their relative positions, z-scores will remain unchanged. Test scores will be examined in their original metric to determine whether they were rising or falling from 1999 to 2001.

[Table 2 About Here]

Table 2 displays the state z-score means of charter schools open in 1999. Several schools did not have three years of test data in state databases. The schools are sorted by whether one, two, or three years of data were available. I decided to omit schools from further analysis if they were missing test scores. Most of the schools with only one or two years of data serve a limited number of students and grade levels. For unknown reasons, their scores also look systematically different from and substantially lower than the scores of three-year schools. About half of the charters are extremely small, with fewer than fifteen students per grade level. States routinely do not report test scores for schools testing fewer than ten students. Many charters bounce above and below this threshold. Another reason for missing data is that charters may start with a few grade levels and then add a grade each year as the oldest cohort moves through the school. A school that opened as a K-3 in 1999, for example, may have added a fourth grade in the school's

second year, when the oldest cohort entered fourth grade. The school's first reported test score would have been in 2000. I estimate that a mismatch on grade levels account for about another 20% of schools with missing data.

For approximately 30% of the schools, I could not determine why test data were missing. Charter school principals may have ignored the state's testing mandate, believing that charters are exempt. Some may give a test other than the one given by the state, indeed, a specific test may have been stipulated in the school's charter. Others may have not wanted to give any test.[17] A few may have closed. Without knowing the reasons for missing scores, the prudent strategy for evaluating achievement from 1999 to 2001 is to examine only those schools reporting test scores for all three years.

Another data challenge pertains to the background characteristics of students, in particular, their socioeconomic status (SES). In order to compare schools fairly, analysts employ statistical controls for differences in student background. The percentage of students on free or reduced lunch is routinely used to model school SES. In the case of charter schools, this statistic is almost certainly inappropriate.[18] A large number of charters do not participate in the free lunch program. Many charters do not have the facilities to offer a hot lunch. Others may have concluded that the federal program is not worth the time and expense. The upshot is that databases reporting students on free lunch, such as the Common Core of Data, underestimate the percentage of poor children attending charters.

In a study involving charter schools in Arizona and Michigan, Hoxby dealt with this problem by using statistics on racial composition to model students' family background. A recent federal study of charter schools, led by RPP International, created its own data set. Researchers surveyed charter school principals and asked them to estimate the percentage of students qualifying for free and reduced lunch. I chose to compute alternative estimates of SES and select the most reasonable one to employ--along with variables for racial composition--in adjusting schools' test scores. The alternative SES estimates make more sense when explained in the context of their effect on achievement scores, so I discuss them below.[19]

A fourth problem arises from the many charter schools serving at-risk students. Charter schools are often created with the purpose of educating students who have not fared well in regular public schools. In other words, these charters explicitly seek and attract students who score very low on achievement tests. Statistically controlling charter schools' demographic characteristics only addresses the bias introduced by this practice to the limited extent that a school's "at-risk" population correlates with demographic variables. Related to this issue are recent studies in both California and Texas that have found large positive effects for charters serving at-risk students. The schools registered significant gains in achievement, outstripping the average gains of all charters in the two states. Charters targeting at risk students may function differently than other charters. It appears wise to identify schools targeting at-risk students for special consideration. I used the Center for Education Reform's *National Charter School Directory, 2001-2002* to code whether charter schools' mission statements explicitly indicate that the schools are seeking an at-risk population. In cases where the school's mission statement was suggestive but unclear, state websites and charter resource sites were consulted.[20]

Another potential source of bias pertains to new schools. Evaluators have noted that charter schools report low test scores for the first year they are open, then recover in subsequent years.[21] This pattern may be due to the stress and distractions educators face in opening a new school, the depressing effect of students moving to new schools, or the propensity of academically struggling students to enter charters. If new schools' scores are initially unrepresentative of school performance over time, the mean achievement for charters in any particular state will be skewed downward for a year in which large numbers of new schools open their doors. And the same phenomenon could inflate computation of charters' value added if first year test scores are used as baseline measures. In the current study, charters that opened in 1999—the first year of the three-year period in which the study evaluates achievement--were coded as new schools, and I pull them out for separate scrutiny.

A sixth consideration relates to school size. Tom Kane and Doug Staiger reviewed several years of test scores in California and North Carolina and found that volatility in test

scores can seriously undermine state accountability systems.[22] Schools frequently deserve a sanction or reward one year--but not the next. Scores in small schools are especially erratic since a few students can drive test scores up or down dramatically. This is a serious concern for studies of charter schools because charters are much smaller than the average public school. In 1999, the median enrollment in charter schools was 137 students, less than one-third of the 475 students enrolled in the median public school.[23] Averaging achievement over three years and two subjects will dampen much of the volatility in test scores. When computing average test scores for charters in each state or for the entire sample, I weighted schools by enrollment. This allows the averages to reflect the number of students that achievement scores actually represent. It also downgrades the weight of small schools in the analysis, however, statistically devaluing a feature that many people believe makes charter schools uniquely attractive.[24]

Two forms of accountability are modeled, the policy instruments used by authorizers and states. I constructed variables representing the accountability of charters to their authorizing agencies from the RPP survey. School principals were asked whether particular facets of their schools—achievement, instruction, and compliance with regulations—were monitored by external authorities. The percentage of principals replying in the affirmative constitutes the authorizer variables, along with the percent of charters closed in a state as an indicator of authorizers' enforcement powers. For state accountability, I created dummy variables reflecting state policies reported in Education Week's *Quality Counts 2001*. The variables reflect whether states rate, reward, or sanction schools based on test scores, and whether they require students to pass an exit exam to receive a high school diploma.[25]

Analysis of Achievement

As raw test scores, the data in Table 3 represent how a reader would view charter schools after examining state test results in the morning newspaper, then comparing charters to other schools in the same state. The z-scores are strikingly low. Charters in these ten states score about one-half standard deviation below average (z-score of -.47). This is approximately equal

to the 32nd percentile, meaning that an astute newspaper reader would discover that 68% of schools score higher than the average charter school. Performance varies by state. Charters in FL, MI, MN, PA, and TX score approximately one standard deviation or more below state means, deficits that are statistically significant. A z-score of –1.00 is equal to the 16th percentile. Charters in CO score well above average (+.44), also statistically significant. In AZ and CA, students in charter schools are achieving at levels comparable to students in the average public school. Massachusetts charters score about a quarter standard deviation below the MA average.

[Table 3 About Here]

In sum, charters in five of the ten states are achieving at statistically significant low levels, well below state averages. Only in Colorado are charter school test scores significantly higher than in regular public schools. It bears repeating, however, that the scores in Table 3 are unadjusted for student background.

[Table 4 About Here]

Table 4 displays the percentage of students eligible for the free lunch program. The statistics for regular public schools, as reported in the Common Core, are located on the left side of the table. In the average public school in the study's ten states, about 34% of students qualify for free lunch. On the right side of the table, four different estimates of SES are presented for charter schools. The CC estimate for the charters in the study is 26%. As discussed above, the Common Core (CC) probably underestimates the percentage of poor children in charter schools.

I computed an alternative estimate, Zip, by matching charters to regular public schools with the same zip code, then assigning the public school data on free lunch (from the CC) to the charter. If more than one public school matched on zip code, I assigned the mean statistic for public schools. Charters without a zip match retained their original CC statistic. The assumption

is that schools with the same zip code, on average, draw students from similar households. The Zip figure of 35% is probably a better estimate of poverty in charters than the CC, but the CC may be a better estimator for large charter schools with significant numbers of poor children. They are the charters most likely to participate in the free lunch program and to collect qualifying information from students' parents. With this in mind, I created a second alternative estimate, Max, by taking the larger of the Zip and CC statistics for each school. The Max variable raises the estimate of poverty in charters to 38%. This is fairly close to the 41% reported in RPP International's 1999 survey of charter school principals in the study's ten states. The fact that Zip, Max, and RPP all converge within a range of six percentage points is somewhat comforting. But the fact remains that the "real" number of poor children in charter schools cannot be determined from these data. The convergence may be misleading.

[Table 5 About Here]

Table 5 shows the impact of different SES variables on regression-adjusted achievement scores. The table displays adjusted z-scores after school SES and racial composition—the percentage of African Americans, Hispanics, and Asians--have been controlled. In other words, the adjusted z-scores compare charter schools to regular public schools serving students with similar racial and socioeconomic backgrounds. Scores are adjusted using the four different measures of SES. As one would expect, the Common Core produces an estimate significantly lower than the three alternative indicators (-.42 vs. alternative estimates of -.29, -.24, and -.26).[26] The three alternative measures suggest as much as one-half of charter schools' initial test score deficit of .47 z-scores may be explained by the demographic characteristics of students entering charter schools. However, charter school achievement still remains about one-fourth standard deviation below that of the average public school after these characteristics are statistically controlled.

Table 5 also shows how pivotal the selection of SES statistic is for evaluating charter achievement in the states. When the Common Core is used to estimate SES, seven states have statistically significant negative scores, only four states using Max, and three using the RPP survey. In PA, charters appear well below average using the Common Core (-.62) but above average (+.29) using the RPP survey. Even the estimates produced by Max and RPP vary by state, although the mean estimates for the entire sample of schools are quite close. The two estimates for charter school achievement in MA (-.53 for Max, +.03 for RPP) and MN (-.44 for Max, -.08 for RPP) differ by more than one-third standard deviation. I will use Max to model SES in the remainder of the paper's analyses, but readers are warned that employing other measures of SES would substantially affect the findings.

[Table 6 About Here]

From 1999 to 2001, charter schools in these tens states notched achievement scores that were significantly below regular public schools. But were charters gaining, losing ground, or staying about the same? In Table 6, test scores from 1999 are presented in their original metric, along with gains and losses registered from 1999 to 2001.[27] Schools in all states reported healthy gains, with regular public schools in Wisconsin the single exception. In eight of the ten states, gains for charter schools were significantly larger than for regular public schools. In the two states where charters gained less, AZ and CA, they scored about the same as public schools in 1999, the baseline year. In CO, charters were higher achieving than public schools in 1999, and they added to that advantage over the next two years.

Although impressive, these gains should be viewed skeptically. Gain scores are erratic. Had the study examined gains from 1999 to 2000, for example, a completely different picture would have emerged (see the z-scores by year in the next display, Table 7). In 2000, regular public schools scored test gains that outpaced those of charter schools. The gains from 1999 to 2001 are almost completely due to an upward spike in charters' scores in 2001. Several

prominent researchers have urged caution when using gain scores to evaluate school performance. Most of the analyses in this paper focus on level of achievement, not gains, and on three-year averages of several tests instead of differences between two years' scores.[28]

[Table 7 About Here]

There is another reason to break out scores by year. I noted above that past evaluations have detected a "first year effect," in which charter schools turn in depressed performances in their first year of operation and rebound in the second year. I compared the scores of charters that were new in 1999 with charters that had been open prior to that year. For the entire period of the study, 1999 to 2001, the new schools' achievement scores were significantly below those of the schools already operating in 1999, with z-scores of -.36 vs. -.20 ($p<.05$, not shown in the table). To probe further, I compared the annual scores for the two groups. Table 7 displays the results. A negative new school effect is evident, but it covers not just the first year but the first two years that charters take students. In 1999, the new schools started -.24 z-scores behind existing charters (-.44 for new schools vs. -.20 for the older schools). They remained approximately the same distance behind in 2000 (-.51 and -.23). And they finally caught up to the older charters in 2001 (z-scores of -.21 and -.23). By 2001, charter schools that were founded in 1999 and before 1999 look almost the same on achievement tests.

What explains the new school effect? The study's data only allow for speculation. As mentioned above, scores may be initially depressed as students and teachers settle into a new school. After two years, achievement returns to its natural level. New charters gain about .30 z-scores in the third year, from 2000 to 2001, which is about nine percentile points around a test's mean. An explanation that is positive for charter schools is also plausible. Charters apparently attract low achieving students. They initially score below students with similar demographic characteristics—a tendency which carries over to school-level scores--but after attending charters for two years, the charter students register learning gains. And the gains exceed those of students

with similar characteristics. This interpretation could be tested by examining test scores of students in the years prior to transferring from a public school to a charter. But what about older schools? Older charters have flat scores of -.20, -.23, -.23, suggesting that gains attributable to charter schooling dissipate sometime after the third year. There is also the possibility that charter competition stimulates better performance from the public schools. If the competitive effects of charters kick in around the third year, rising scores for both charters and regular pubic schools would produce flat charter z-scores after the third year. Examining regular public schools' scores before and after charters are established in the same neighborhood would be revealing on that question.

I discuss implications of the study's findings at the conclusion of the paper, but two are worth pointing out now. First, the two-year new school effect is a constraint that state officials should recognize in building effective accountability systems. If achievement is depressed from the stress of opening a new charter school, a two-year grace period may be advisable before "starting the clock" on sanctioning new charters. Texas exempts charters' first year of test scores from the state's accountability system. Federal legislation requires states to begin some form of intervention after three straight years of failing tests scores. New charter schools are especially vulnerable to this penalty. The first two years that a charter is open should be treated differently. Otherwise, on top of everything else, schools struggling to get up and running will be prematurely placed on watch lists before test score gains have had a chance to materialize.[29] A second implication is that accountability systems based on gain scores will generate inflated gains for charter schools if the first year or two of test scores are artificially low and used as baselines. This touches upon the question of how gain scores can be appropriately used in accountability systems, an issue relevant to all schools but one with special salience for charter schools.

[Table 8 About Here]

Table 8 examines whether charters targeting at risk students are unduly skewing the overall appearance of charters on state tests. Indeed, the forty-nine charters with an at-risk focus score significantly below other charters (-1.05 vs. -.17) . Removing them from the sample improves charters' test scores by .07 z-scores (from -.24 for all charters to the -.17 reported here for non at-risk charters). Controlling for SES and racial composition only explains about one-third of the at-risk schools' achievement deficit. The characteristics which place students at-risk are far more complex than demographic statistics can capture.

This topic highlights the difficulty of comparing school achievement. What is the fairest way of comparing charter schools to regular public schools? Discounting the results from at-risk charters may leave a pool of charters that are different from public schools in such fundamental ways that the comparison is artificial. Mainstream public schools may not seek out and select at-risk students, but they serve thousands of children for whom learning is a struggle. Unless at-risk students are enrolled in special education or limited English speaking programs, their test scores count in the reporting of public schools' scores. At the same time, however, it is difficult to say what a fair comparison group would be for charters—or any other schools—that primarily enroll at-risk students. The same problem crops up for continuation schools or alternative settings for students who have been expelled or have temporarily dropped out of school. The question affects much more than academic research. In offering rewards and imposing sanctions, state officials must decide how to treat schools fairly when the schools' primary mission is educating students who are extremely difficult to educate.

As shown in Table 8, many of the individual state counts of at-risk schools are small, so their impact on state means is muted. The effect of the at-risk targets on achievement appears greatest in two states, FL and TX. In Florida, eleven charter schools in the study target at risk students, eighteen do not. The adjusted score of at-risk charters is –1.47, placing them among the bottom 7% of schools in the state. For non-at-risk charters, the adjusted z-score is 0.03, about the state average. To return to the reader who compares the test scores of schools as published in the morning newspaper, the initial impression in Florida would be that charter

schools are doing quite poorly (-0.98 z-score).[1] If the scores are adjusted to control for SES and racial composition, the performance of Florida's charters will seem somewhat better, but still below average (-0.37). If only charters with a general clientele are considered, omitting schools with an at-risk specialty, the state's charters score about average (0.03).

In Texas, a strange pattern appears. The schools targeting at-risk students test higher than the non-at-risk schools. The at-risk schools' raw scores are better (–1.11 vs. –1.76), and after employing demographic controls, the gap widens (-.38 vs. –1.37)[30]. The finding is based on only six schools targeting at-risk students and must be taken with a grain of salt, However, it mirrors a finding in the most recent evaluation of Texas charter schools conducted by a consortium of Texas research centers.[31] Texas is unique in legislating a separate category for charters serving at risk students, known as 75 Percent Rule charter schools. These schools must enroll at least 75% at-risk students.[32] It is puzzling that after combing through state records, the Texas researchers discovered that many of the 75% rule charters report absolutely no at-risk students attending their schools. And about one-third of the general open-enrollment schools report more than 75% of students at-risk. The demographic data from charter schools are undoubtedly flawed.

So what should be made of the performance of Texas charters? Overall, they achieve at extremely low levels. But, as indicated above, they are making significant gains. Notwithstanding deficiencies in the data, something positive may be happening in Texas's at-risk charters. In 1999, only a small number of at-risk charters had been established, which explains why so few show up in the current study. It is unwise to draw firm policy guidance until the data problems are solved and a more thorough evaluation is conducted. However, Texas ultimately may provide examples of charter schools working successfully with students who traditionally underachieve on tests of academic performance.

[Table 9 About Here]

Table 9 breaks out scores by grade level and school subjects, comparing achievement in the 4th, 8th, and 10th grades and in reading and math. A school may be counted in two of the table's grade categories. A K-8 school, for example, contributes scores from both 4th and 8th grades. The big story emerging from Table 9 is the strength of the reading scores. They outstrip math scores at all three grade levels. In eighth and tenth grade reading, charter schools perform about the same as the average public school. In both subjects, it is also notable that scores are highest in eighth grade.

Speculation is necessary here. These grade and subject patterns diverge from the gains detected by national and international tests of U.S. academic achievement. Math gains have consistently trumped those in reading. On the NAEP trend test, reading is a subject in which public schools have not shown much progress since the 1980s. Math scores have risen slowly. On the main NAEP test, dramatic gains in mathematics in the 1990s outpaced flat reading scores. As for the performance of different grade levels, American elementary students consistently do better than older American students on international tests. In mathematics and science, the U.S. scores highest at the fourth grade, slumps at the 8th grade, and performs even worse on tests in the 12th grade. Charters may be displaying relative strength in reading compared to math--and in the middle/high school grades compared to 4th grade--because public schools are weakest in this subject and these grades. Like a small fish that has been moved to a smaller pond, charter school achievement may appear larger in a different context although its true size remains unchanged.

Analysis of Accountability

Charter schools are held accountable by two authorities. Authorizing agencies grant charters and monitor charter school operations. State assessment and accountability systems measure student achievement in all schools, including charters, and implement rewards and sanctions based on the results. In Massachusetts, the state is the sole authorizer of new charter schools.[33] Separate units within the Massachusetts Department of Education administer the

charter school program and the state's assessment and accountability system. In states such as FL and PA, authorizers are primarily local school boards. In MI and AZ, charters are granted by a wide range of authorizers, including universities and other public entities. In several states, charters may appeal local rejection of a charter petition to state officials.

[Table 10 About Here]

Table 10 examines the accountability tools of authorizers. By focusing on policies rather than schools, the study's N shrinks from a few hundred schools to ten states as all schools in the same state share the same policy regime. Any trends in the data must be viewed cautiously. I computed simple correlation coefficients for charter school achievement (with scores adjusted for student background) and external monitoring of charters. As explained above, the RPP survey asked charter school principals whether several school functions are monitored. Although the survey did not identify a specific external authority, in most of states, chartering agencies are the legal monitors of charters' activities. I also included the percentage of charter schools that have been closed down by authorizers, a statistic reflecting authorizers' willingness to implement sanctions. In Table 10, the only variable approaching a significant correlation with charter achievement is the percentage of closures, which is negatively related ($p=.07$). Low scoring charters (relative to public schools) tend to be located in states where a large number of charters have been shut down.

[Table 11 About Here]

Not too much should be made of this. It may mean that promiscuously closing down charters depresses the performance of charters that remain. But a more likely explanation is that the correlation shows that authorizers are doing their jobs, stepping in where charter schools are not functioning very well.[34] Table 11 looks at closures more closely. In the table, states are

ordered by closure rate. The negative correlation with achievement is quite clear. States with lower scoring charters are at the top of the list, states with higher scoring charters are at the bottom. But also notice that two states stand out with the highest closure rates, MN and TX. They have aggressively opened charters serving low achieving students. In Table 11, the "at-risk" statistic reports the percentage of charters in the state targeting at-risk students. The variable "relative poverty" shows the difference between charter schools and regular public schools on students eligible for free lunch. In Minnesota, charters average 22% points of more free lunch students than the average traditional public school. Minnesota and Texas are the only two states with above average at-risk and relative poverty statistics. Their closed charters were located in poor communities in which suitable facilities, well-trained teachers, and investment capital for new schools are scarce. These schools failed to offer the education that they promised, and authorizers shut them down. Also note that the underlying numbers are small. Minnesota's closure rate of 8% represents six charter schools that were shut down out of seventy-four in the state.[35]

[Table 12 About Here]

Table 12 examines the relationship of charter achievement and state accountability policies. Three policies apply to schools—whether the state releases annual ratings of schools, offers a reward, or imposes sanctions. One policy is aimed at students, whether they must pass an exit exam before receiving a high school diploma. All of the relationships are negative. Sanctions are statistically significant, and the negative effect of exit tests nearly reaches significance.

Tables 11 and 12 report similar findings for both authorizer and state accountability. Punitive policies—closing charters, invoking school sanctions, requiring student exit tests--are correlated with negative achievement in charter schools. Correlation is not causality. The intervention may or may not cause the low achievement. Authorities intervene—closing

charters, applying sanctions, requiring exit tests—when test scores are low. I ran the same correlation coefficients on the charter school gains from 1999 to 2001, first converting the gains to standard deviations of 1999 baseline scores. If the policies do indeed depress achievement, they should show a negative correlation with achievement gains as well. The results are displayed in Tables 13 and 14. Most of the relationships continue to be negative, but none reach statistical significance. External monitoring of compliance with regulations switches signs, from negative to positive (see Table 13), as does states' rankings of schools on academic performance (see Table 14).

[Tables 13 & 14 About Here]

What could explain punitive policies exerting a negative effect on charter schools net of their effect on public schools? State penalties might be influencing the type of student that seeks charter schooling. Compared to parents in states without exit exams, parents of low performing students in states with exit exams may be more motivated to take a chance on charter schools. Regular public schools that have been sanctioned by the state may also lose students with low test scores who seek a better education. This is speculation, of course, but the demand for charters among the families of low achieving students may be greater in states where low achievement is penalized, regardless of whether the penalties fall on students or schools. If such families leave public schools for new charter schools, the test scores of public schools would look better—at least temporarily--and the scores of charters would start out depressed.

A second possibility is related to the regulatory environment. Strict oversight by charter authorizers, sanctions on low performing schools, and mandatory student exit tests may be accoutrements of top-down, bureaucratic school systems. A large portion of charter school authorizers are local school districts. If the local political culture encourages close monitoring and tight regulation, accountability systems could furnish the system with several tools to promote standardization.[36] Charter schools would suffer. As charter schools mature in dozens of

states, comparisons of state and local political cultures will allow more thorough investigations into the interaction of charter school performance and local regulatory environments.

Summary and Discussion

This study analyzed academic achievement in the charter schools of ten states from 1999 to 2001. Achievement is significantly lower in charters than in regular public schools, by about one-half standard deviation on raw test scores and one-fourth standard deviation when adjusted for students' racial and socioeconomic backgrounds. Charters serving at-risk students achieve at significantly lower levels than open admission charters. New charter schools achieve at lower levels than existing charters for the first two years, but catch up with older charters by the third year. Charter achievement is stronger in reading than in math and stronger in eighth grade than in fourth and tenth grades. Charters were apparently narrowing the gap during this period by producing larger learning gains than public schools.

This study's principal findings are negative for charter schools. On tests of academic achievement, charter schools in the study scored significantly lower than regular public schools with similar students. I began by pointing out several technical decisions concerning the data and how they would be analyzed. Most of the decisions—omitting schools with missing data, computing an alternative measure of SES, weighting scores by enrollment—benefited charter schools' test scores. If I had treated the data differently, then charter test scores would have appeared even lower than depicted here. Moreover, I analyzed performance from several different perspectives—taking into consideration whether charters target at-risk students, whether the schools were just getting started, and how they performed in different subjects and different grades. Some of these perspectives discovered charter achievement to be statistically indistinguishable from achievement in regular public schools. Not one uncovered a statistically significant charter school advantage.

The study uncovered no evidence that accountability policies produce higher achievement in charter schools. Three policies are negatively related to charter schools test

scores--closing charters, imposing school sanctions, and requiring students to pass exit tests. The data do not allow for causal claims. I proposed two possible explanations and discussed how they could be investigated. The fact that policies with a punitive cast are correlated with low achievement in charter schools suggests that low achieving students may be more likely to seek out charter schooling in high stakes settings. Careful study of charter school enrollment patterns in high and low stake states may be able to untangle cause from effect on this topic. Accountability policies also may be related to states' regulatory environments. Studies with a large sample of states might be able to detect whether the amount of state regulation is related to charter school performance.

Three important considerations temper the pessimism of the findings. The test score gains of charter schools from 1999 to 2001 indicate that low achievement in charters may be a temporary condition. Charters may attract students who are low achievers before they ever set foot on a charter school campus. A rigorous analysis of gain scores would be better able to tease out any selection effects from charter schools enrolling students who are, net of the controls employed here, low achievers. If charter schools represent a remedial policy intervention—a policy designed to tackle poor academic performance--then initial low test scores are not surprising. Analysis of student-level data would also help to isolate the effect of charter schools.

A second consideration is that competition from charters may raise achievement in both charters and regular public schools. This is certainly suggested when gain scores from 1999 to 2001 are examined, but it is called into question by the instability of the gains. As noted above, Hoxby discovered positive effects for charters on public school test scores in Arizona and Michigan, two states in the current study.[37] Whether "all boats are rising" needs to be explored in additional states.

The third point relates to productivity. In seven of the study's ten states, open admission charters, schools that do not target at-risk students, produce test scores that are statistically indistinguishable from the scores of regular public schools. Charter schools are less

costly to operate than regular public schools. If charters can produce the same amount of learning as the average public school, that is a strong argument in charters' favor.

The study uncovers two policy areas for states to address in order to integrate charters fully and fairly into accountability systems. Better data must be collected from charter schools, including annual test scores and reliable demographic information. The federal No Child Left Behind Act mandates annual achievement testing in grades three through eight, so the situation is likely to improve soon. Better information will not only serve the cause of accountability. It will allow parents who are considering charter schooling to make more informed decisions. States should also consider a two-year moratorium on the "accountability clock" for new charter schools. If test scores start out extremely low, charter schools may face sanctions before they have had a real chance to prove themselves.

Finally, researchers and policymakers should be watchful for how accountability systems and charter schools interact in the future. I began the paper by observing that charter schools are a product of both the choice and accountability movements. Charters offer choice advocates an alternative to traditional public schools run by local public districts. They offer accountability advocates the means to close down charters that persistently fail. There is, however, the seed of a conflict between the two ideas. Accountability systems are built on the premise that schools must raise academic achievement. A conflict looms when parents value school attributes that are not valued in accountability systems. If scores on achievement tests are more important to the state than they are to parents, accountability systems will impose penalties on schools despite the schools' popularity with parents. In the future, if this fundamental conflict surfaces very often, the politics of choice and accountability will undergo a profound, perhaps wrenching change.

Table 1 Obstacles Encountered While Attempting to Analyze Student Achievement in Charter Schools, Ten States, 1999-2001

Problem	Strategy	Limits of Strategy
Tests--Non-standard measures. Different content, test protocols, and scoring.	Horse-race schools within states by computing z-scores	Portrays tests that are incomparable as comparable. Can't capture "all boats rising."
Missing test data for charters, mainly due to size and grades open.	Restrict study to schools with complete data	Sharply reduced N & dropped schools look systematically different. Could positively skew charter scores.
SES—faulty free and reduced lunch estimates in Common Core. AZ does not report.	Compute alternative estimates of SES and look for convergence	No polar star—don't know the "real" statistic.
School size and score volatility	Scores averaged over three years, weighted by enrollment.	Effect of small schools diminished.
Targeting at risk students-- In TX, "75%" schools.	Code from school's mission statement in CER directory.	May bias the charter sample, make non-comparable to public schools.

Table 2 Charter Schools, Achievement Z-Scores, 1999-2001
(N= 638)

State	Schools w/3Yrs of Data	Schools w/2Yrs of Data	Schools w/1Yr of Data
AZ (N=140)	+0.01 (n=56)	-0.56* (n=55)	-0.09 (n=29)
CA (N=122)	0.00 (n=98)	-0.43* (n=20)	-0.65* (n=4)
CO (N=44)	+0.44* (n=31)	+0.01 (n=12)	+0.09 (n=1)
FL (N=54)	-0.98* (n=29)	-0.40 (n=19)	+0.04 (n=6)
MA (N=31)	-0.26 (n=21)	-0.80* (n=9)	+1.38 (n=1)
MI (N=116)	-1.06* (n=84)	-1.17* (n=18)	-1.08* (n=14)
MN (N=24)	-1.22* (n=16)	-1.07 (n=5)	-1.13 (n=3)
PA (N=27)	-1.14* (n=11)	-1.74* (n=12)	-1.73* (n=4)
TX (N=57)	-1.20* (n=25)	-2.33* (n=22)	-1.90* (n=10)
WI (N=23)	-0.64 (n=11)	-0.67 (n=5)	-0.74 (n=7)
Average (N=638)	-0.47* (N=382)	-0.88* (N=177)	-0.67* (N=79)

* $\rho < .05$, two-tailed test of z-score = 0

Table 3 Achievement Scores of Charter Schools in the Study, 1999-2001 (Within-state Z-scores, weighted by enrollment)

State	Raw Z
AZ	+0.01 (n=56)
CA	0.00 (n=98)
CO	+0.44* (n=31)
FL	-0.98* (n=29)
MA	-0.26 (n=21)
MI	-1.06* (n=84)
MN	-1.22* (n=16)
PA	-1.14* (n=11)
TX	-1.20* (n=25)
WI	-0.64 (n=11)
Average	-0.47* (N=382)

* $\rho < .05$, two-tailed test of z-score = 0

Table 4 Free Lunch Estimates
(Mean % in schools)

	Regular	Charter			
State	CC	CC	Zip	Max	1999 RPP
AZ	43% (n=814)	56% (n=18)	40% (n=51)	42% (n=51)	43% (n=155)
CA	41% (n=6,655)	27% (n=92)	36% (n=98)	40% (n=98)	31% (n=143)
CO	26% (n=766)	10% (n=31)	18% (n=31)	18% (N=31)	18% (n=57)
FL	41% (n=2,348)	32% (n=29)	40% (n=29)	42% (n=29)	42% (n=60)
MA	24% (n=1,257)	0% (n=21)	17% (n=21)	17% (n=21)	37% (n=32)
MI	30% (n=2,351)	27% (n=83)	37% (n=84)	43% (n=84)	39% (n=121)
MN	22% (n=1,169)	41% (n=16)	35% (n=16)	44% (n=16)	60% (n=37)
PA	24% (n=2,588)	0% (n=11)	45% (n=11)	45% (n=11)	64% (n=22)
TX	39% (n=5,660)	33% (n=25)	44% (n=25)	48% (n=25)	62% (n=72)
WI	18% (n=1,833)	19% (n=11)	21% (n=11)	21% (n=11)	28% (n=26)
Average	34% (N=25,441)	26% (N=337)	35% (N=377)	38% (N=377)	41% (N=725)

Note: CC statistics are the percent of students on free lunch as reported in the Common Core of Data, collected annually by the National Center for Education Statistics. The three alternative SES measures were computed for charter schools only. ZIP is the mean percentage of students on free lunch in public schools with the same zip code. MAX is the larger of ZIP and CC for each charter. RPP is the percentage of students eligible for free and reduced lunch as reported in a survey of charter school principals conducted by RPP in 1999.

Table 5 Mean Achievement Adjusted by SES Measures

(Within-state Z-scores for charter schools, adjusted for SES and racial composition, weighted by enrollment)

State	CC	Zip	Max	RPP
AZ	-0.23 (n=18)	-0.06 (n=51)	-0.03 (n=51)	+0.01 (n=56)
CA	-0.16* (n=92)	-0.09 (n=97)	-0.02 (n=97)	-0.19* (n=97)
CO	+0.01 (n=31)	+0.18 (n=31)	+0.18 (n=31)	+0.13 (n=31)
FL	-0.60* (n=29)	-0.39 (n=29)	-0.37 (n=29)	-0.39 (n=29)
MA	-0.81* (n=21)	-0.52* (n=21)	-0.53* (n=21)	+0.03 (n=21)
MI	-0.74* (n=83)	-0.67* (n=84)	-0.63* (n=84)	-0.66* (n=84)
MN	-0.51* (n=16)	-0.72* (n=16)	-0.44* (n=16)	-0.08 (n=16)
PA	-0.62* (n=11)	+0.05 (n=11)	+0.05 (n=11)	+0.29 (n=11)
TX	-1.17* (n=25)	-1.11* (n=25)	-1.09* (n=25)	-1.03* (n=25)
WI	-0.24 (n=11)	-0.18 (n=11)	-0.18 (n=11)	+0.01 (n=11)
Average	-0.42* (N=337)	-0.29* (N=376)	-0.24* (N=376)	-0.26* (N=381)

* $\rho < .05$, two-tailed test of z-score = 0

Note: CC statistics are the percent of students on free lunch as reported in the Common Core of Data, collected annually by the National Center for Education Statistics. The three alternative SES measures were assigned to charter schools only. ZIP is the mean percentage of students on free lunch in public schools with the same zip code. MAX is the larger of ZIP and CC for each charter. RPP is based on the percentage of students eligible for free and reduced lunch as reported in a survey of charter school principals conducted by RPP in 1999. School level data from the RPP survey were unavailable. Each state's mean was assigned to all charter schools in the state.

Table 6 Adjusted Achievement Gains of Charter Schools and Regular Public Schools, 1999-2001
(Test scores adjusted for SES and racial composition, weighted by enrollment)

State	Test	Metric	1999 Score		Mean Gain Score (SD)	Gain 1999-2001	
			Charter	Regular		Charter	Regular
AZ	SAT-9	Percentile	50.5	51.4	+1.5 (0.6)	+0.3	+1.5*
CA	SAT-9	Percentile	44.9	44.4	+5.9 (1.0)	+4.1	+5.9*
CO	CSAP	% Proficient	60.3	51.8	+5.0 (0.4)	+5.9*	+5.0
FL	FCAT	Scale Score	266.5	293.8	+9.3 (1.3)	+13.2*	+9.3
MA	MCAS	Scale Score	230.6	233.0	+4.2 (0.3)	+4.6*	+4.2
MI	MEAP	% Satisfactory	40.4	62.3	+2.5 (1.2)	+3.3*	+2.5
MN	BST/ MCA	Scale Score	825.4	981.1	+34.8 (17.6)	+54.1*	+34.7
PA	PSAT	Scale Score	1156.6	1305.2	+5.6 (3.8)	+27.0*	+5.6
TX	TAAS	% Passing	73.2	87.5	+3.9 (1.7)	+7.1*	+3.9
WI	CTBS-5	Percentile	57.5	66.6	-1.0 (0.7)	+2.7*	-1.0

* $\rho < .05$, two-tailed test of Charter = Regular

Table 7 Annual Adjusted Achievement of Charters Opening in 1999
(Within-state Z-scores, adjusted for SES and racial composition, weighted by enrollment, compared to schools already open in 1999)

	1999		2000		2001	
State	New	Existing	New	Existing	New	Existing
AZ	-0.04 (n=6)	+0.04 (n=45)	+0.16 (n=6)	-0.06 (n=45)	+0.18 (n=6)	-0.18 (n=45)
CA	+0.03 (n=15)	-0.02 (n=82)	-0.06 (n=15)	-0.03 (n=82)	-0.09 (n=15)	-0.02 (n=82)
CO	-0.27 (n=4)	+0.25 (n=27)	-0.23 (n=4)	+0.07 (n=27)	+0.56 (n=4)	+0.20 (n=27)
FL	-0.25 (n=19)	-1.12* (n=10)	-0.19 (n=19)	-1.34 (n=10)	-0.05[a] (n=19)	-1.08* (n=10)
MA	-0.53 (n=3)	-0.51* (n=18)	-0.63 (n=3)	-0.51* (n=18)	-0.67 (n=3)	-0.84* (n=18)
MI	-0.89* (n=20)	-0.58* (n=64)	-1.18* (n=20)	-0.70*[a] (n=64)	-0.31 (n=20)	-0.56* (n=64)
MN	-0.02 (n=3)	-0.46 (n=13)	-0.18 (n=3)	-0.56* (n=13)	-0.76 (n=3)	-0.41 (n=13)
PA	-0.76* (n=8)	+0.45[a] (n=3)	-0.18 (n=8)	+0.14 (n=3)	+0.08 (n=8)	+0.70 (n=3)
TX	-1.08* (n=15)	-1.55* (n=8)	-1.30 (n=15)	-0.39 (n=8)	-0.87 (n=15)	-1.46* (n=8)
WI	-3.02 (n=1)	-0.16 (n=10)	-1.73 (n=1)	-0.16 (n=10)	-2.45 (n=1)	-0.16 (n=10)
Average	-0.44* (n=94)	-0.20*[a] (n=280)	-0.51* (n=94)	-0.23*[a] (n=280)	-0.21 (n=94)	-0.23* (n=280)
All Charters	-.26* (n=376)		-.31* (n=376)		-.23* (n=376)	

* $\rho < .05$, two-tailed test of z-score = 0

a $\rho < .05$, two-tailed test of New = Existing

Note: New charter schools opened between July, 1998 and June, 1999.
Existing charters were open before July, 1998.

Table 8 Adjusted Achievement of Charter Schools Targeting At-Risk Students (Within-state Z-scores, adjusted for SES and racial composition, weighted by enrollment)

	At-Risk		Non-At-Risk	
State	Raw Z	Adjusted Z	Raw Z	Adjusted Z
AZ	-1.40* (n=5)	-0.81* (n=5)	+0.14 (n=51)	+0.02 (n=46)
CA	-1.21* (n=15)	-1.19* (n=15)	+0.03 (n=83)	+0.06 (n=82)
CO	NA (n=0)	NA (n=0)	+0.31 (n=31)	+0.18 (n=31)
FL	-2.03* (n=11)	-1.47* (n=11)	-0.34 (n=18)	+0.03 (n=18)
MA	NA (n=0)	NA (n=0)	-0.51* (n=21)	-0.53* (n=21)
MI	-1.27* (n=4)	-1.28* (n=4)	-1.06* (n=80)	-0.59* (n=80)
MN	-2.26* (n=4)	-1.10* (n=4)	-0.87* (n=12)	-0.25 (n=12)
PA	-2.06 (n=1)	+0.06 (n=1)	-1.14* (n=10)	+0.05 (n=10)
TX	-1.11 (n=6)	-0.38 (n=6)	-1.76* (n=17)	-1.37* (n=17)
WI	-1.77* (n=3)	-1.82* (n=3)	+0.06 (n=8)	+0.22 (n=8)
Average	-1.40* (n=49)	-1.05* (n=49)	-0.30* (n=331)	-0.17* (n=325)

* $\rho < .05$, two-tailed test of z-score = 0

Note: Charter policies were recorded from *The National Charter Schools Directory 2000-2001*, (Center for Education Reform, 2002), and state department websites.

Table 9 Adjusted Achievement of Charter Schools by Grade-Subject

(Within-state Z-scores, adjusted for SES and racial composition, weighted by enrollment)

	4th Grade		8th Grade		10th Grade	
State	Reading	Math	Reading	Math	Reading	Math
AZ	+0.13 (n=38)	+0.05 (n=38)	0.00 (n=40)	-0.12 (n=40)	+0.30[a] (n=19)	-0.08 (n=19)
CA	+0.08 (n=66)	+0.13[a] (n=66)	+0.13[a] (n=54)	-0.26* (n=53)	+0.02[a] (n=30)	-0.42 (n=30)
CO	+0.16 (n=31)	+0.22 (n=31)	+0.25 (n=25)	+0.16 (n=25)	NA	NA
FL	-0.27 (n=12)	+0.01[a] (n=11)	-0.11 (n=17)	-0.19 (n=18)	-0.36 (n=9)	-0.59 (n=9)
MA	-0.94* (n=17)	-0.92* (n=17)	-0.35 (n=18)	-0.53* (n=18)	+0.52 (n=5)	+0.20 (n=5)
MI	-0.76* (n=78)	-0.84* (n=78)	-0.26* (n=65)	-0.33* (n=61)	NA	NA
MN	-0.52* (n=12)	-0.71* (n=13)	-0.23 (n=9)	-0.33 (n=8)	NA	NA
PA	-0.20[a] (n=8)	-0.50 (n=8)	+0.80[a] (n=10)	+0.55 (n=10)	+1.15 (n=3)	+0.69 (n=3)
TX	-0.71[a] (n=18)	-1.47* (n=18)	-0.05 (n=14)	-1.04 (n=14)	-1.35*[a] (n=11)	-1.89* (n=11)
WI	+1.34* (n=3)	+0.96 (n=3)	+0.31 (n=5)	-0.09 (n=5)	-2.20* (n=5)	-2.05* (n=5)
Average	-0.26*[a] (n=283)	-0.40* (n=283)	-0.02[a] (n=254)	-0.24* (n=252)	-0.04[a] (n=82)	-0.44* (n=82)

* $\rho < .05$, two-tailed test of z-score = 0

a $\rho < .05$, two-tailed test of Reading = Math

Note: AZ did not give a high school test in 2001. CO and MN did not give the same high school math and reading tests in all three years, 1999-2001. We were unable to obtain building-level high school test data from MA (2001) and MI (1999-2001) before completion of study. Alternative grades were used in the following states: CO, 5th grade math; MI, 7th grade; MN, 3rd grade; PA, 5th and 11th grades.

Table 10 Charter Achievement and Authorizers' Accountability Mechanisms

Type	Correlation Coefficient	ρ Value
% Closures	-.59	0.07
Achievement	-.23	0.52
Instruction	-.13	0.72
Compliance with Regulations	-.33	0.36

Note: Pearson correlation coefficient.

Table 11 Charter Achievement and State Closure Rates

State	% Closures	Adjusted Z	At-Risk	Relative Poverty
MN	8.1%	-0.44*	25.0%	+22%
TX	5.9%	-1.09*	26.1%	+9%
AZ	4.7%	-0.03	9.8%	-1%
MA	4.7%	-0.53*	NA	-7%
FL	4.4%	-0.37	37.9%	+1%
MI	3.1%	-0.63*	4.8%	+13%
CO	2.4%	+0.18	NA	-8%
WI	2.2%	-0.18	27.3%	+3%
CA	2.1%	-0.02	15.5%	-1%
PA	1.5%	+0.05	9.1%	+21%
Average	3.9%	-0.24*	19.4%	+5%

* $\rho < .05$, two-tailed test of z-score = 0

Note: Closure data provided by the Center for Education Reform (CER). At-Risk is the percentage of charter schools in the study targeting At-Risk students. Relative Poverty is the difference between the percentage of free lunch students in charter schools and regular public schools (reported in Table 4).

Table 12 Charter Achievement and State Accountability Policies

Type of Policy	Policy	No Policy
Rating	-0.34 (n=6)	-0.26 (n=4)
Reward	-0.30 (n=6)	-0.31 (n=4)
Sanction	-0.66 (n=3)	-0.15* (n=7)
Exit Test	-0.63 (n=3)	-0.17 (n=7)

* $\rho < .05$, two-tailed test of Policy = No Policy

Note: States planning on implementing a policy in the future are counted as not having it. The ρ-value of Exit Test is 0.07.

Table 13 Charter Gains and Authorizers' Accountability Mechanisms

Type	Correlation Coefficient	ρ Value
% Closures	-.30	0.39
Achievement	-.27	0.45
Instruction	-.10	0.78
Compliance with Regulations	.49	0.15

Note: Pearson correlation coefficient.

Table 14 Charter Gains and State Accountability Policies

Type of Policy	Policy	No Policy
Rating	+1.98 (n=6)	+1.39 (n=4)
Reward	+1.46 (n=6)	+1.93 (n=4)
Sanction	+1.62 (n=3)	+2.03 (n=7)
Exit Test	+1.64 (n=3)	+1.98 (n=7)

Note: No relationship significant, $\rho < .05$. States planning on implementing a policy in the future are counted as not having it.

Notes

[1] Charter school reform in the 1990s is detailed in Bryan C. Hassel, *The Charter School Challenge*, (Brookings Institution Press, 1999). State accountability systems are described in Margaret E. Goertz and Mark C. Duffy, *Assessment and Accountability Systems in the 50 States: 1999-2000*, CPRE Research Report Series #RR-046, Consortium for Policy Research in Education, Philadelphia, PA, March 2001.

[2] Gary Miron and Christopher Nelson, "Student Academic Achievement in Charter Schools: What We Know and Why We Know So Little," National Center for the Study of Privatization in Education, Teachers College, Columbia University, 2001.

[3] Paul Hill, Robin Lake, Mary Beth Celio, Christine Campbell, Paul Herdman, Katrina Bulkley, *A Study of Charter School Accountability*, (U.S. Department of Education, Office of Educational Research and Improvement, 2001).

[4] Robert V. Antonucci, "1997 Test Results from Massachusetts Charter Schools," Massachusetts Department of Education, 1997; Lori A. Mulholland, "Arizona Charter School Progress Evaluation," Morrison Institute for Public Policy, Arizona State Univeristy, 1999.

[5] Miron and Nelson, 2001.

[6] See also Lew Solomon, K. Paark, and D. Garcia, "Does Charter School Attendance Improve Test Scores? The Arizona Results" Goldwater Institute, Phoenix, AZ, 2001; Eric Bettinger, "The Effect of Charter Schools on Charter Students and Public Schools," National Center for the Study of Privatization in Education, Teachers College, Columbia University, 1999; Randall W. Eberts and Kevin M. Hollenbeck, "An Examination of Student Achievement in Michigan Charter Schools," W.E. Upjohn Institute for Employment Research, Kalamazoo, MI, 2001. Caroline M. Hoxby, "How School Choice Affects the Achievement of Public School Students," Koret Task Force, Hoover Institution, September 2001; Caroline M. Hoxby, "School Choice and School Productivity (or Could School Choice Be A Tide That Lifts All Boats?)" Paper presented at the NBER Conference on the Economics of School Choice, Cambridge, MA, February 23-24, 2001.

[7] Brian P. Gill, P. Michael Timpane, Karen E. Ross, Dominic J. Brewer, *Rhetoric Versus Reality: What We Know and What We Need to Know About Vouchers and Charter Schools*, RAND Corporation, Santa Monica, CA, 2002.

[8] Linda M. McNeil, *Contradictions of School Reform*, (Routledge, 2000); David Lee Stevenson and Kathryn S. Schiller, "State Education Policies and Changing School Practices: Evidence from the National Longitudinal Study of Schools, 1980-1993," *American Journal of Education*, 107 (4) August 1999, pp. 261-288.

[9] Most of the dropout studies examine the effect of minimum competency exams in the 1980s. See James S. Catterall, "Standards and School Dropouts: A National Study of Tests Required for High School Graduation," *American Journal of Education*, 98 (1) November, 1989, pp. 1-34 Brian Jacob, "Getting Tough? The Impact of Mandatory High School Graduation Exams on Student Achievement and Dropout Rates." *Educational Evaluation and Policy Analysis*, 23 (2) Summer 2001, pp. 99-122; Kathryn S. Schiller and Chandra Muller, "External Examinations and Accountability, Educational Expectations, and High School Graduation," *American Journal of Education*, 108 (2) February, 2000, pp.73-102.

[10] John H. Bishop, "The Effect of Curriculum-Based External Exit Exams on Student Achievement," *Journal of Economic Education*, 29 (2) Spring 1998, pp. 171-182.

[11] John H. Bishop, "Nerd Harassment, Incentives, School Priorities, and Learning," in *Earning and Learning*, Susan E. Mayer and Paul E. Peterson, Eds., (Brookings Institution Press, 1999), pp. 231-279.

BEST COPY AVAILABLE

[12] John H. Bishop, "Strengthening Incentives for Student Effort and Learning: Michigan's Merit Award Program" Working Paper 01-10, (Center for Advanced Human Resource Studies, Cornell University, 2001).

[13] John H. Bishop, Ferran Mane, Michael Bishop, and Joan Moriarty, "The Role of End-of-Course Exams and Minimum Competency Exams in Standards-Based Reforms, in *Brookings Papers on Education Policy 2002*, Diane Ravitch, Ed., (Brookings Institution Press, 2002), pp. 267-345. Also see John Bishop, Ferran Mane, and Michael Bishop, "Is Standards-Based Reform Working?...and For Whom?" Working Paper 01-11, (Center for Advanced Human Resource Studies, Cornell University, 2001).

[14] David W. Grissmer, Ann Flanagan, Jennifer Kawata, Stephanie Williamson, *Improving Student Achievement: What State NAEP Test Scores Tell Us*, (RAND Corporation, 2000).

[15] Figlio and Page analyze the distributional consequences for Florida allocating vouchers to students in schools with failing test scores. Figlio, David N., and Marianne E. Page, "Can School Choice and School Accountability Successfully Coexist?" Working Paper, National Bureau of Economics Research, June 2001.

[16] Chester E. Finn, Jr., Bruno V. Manno, and Greg Vanourek, *Charter Schools in Action*, (Princeton University Press, 2000).

[17] Hill, et. al. report that many charter school principals object to state achievement tests (pp. 52-53).

[18] Jerry Horn and Gary Miron, "An Evaluation of the Michigan Charter School Initiative: Performance, Accountability, and Impact," (The Evaluation Center, Western Michigan University, 2000); Caroline M. Hoxby, "How School Choice Affects the Achievement of Public School Students," Prepared for Koret Task Force meeting, The Hoover Institution, Palo Alto, CA, September 20-21, 2001; Randall V. Eberts and Kevin Hollenbeck, "The State of Charter Schools in Colorado 1999-2000," (Colorado State Department of Education, 2001).

[19]RPP International, *The State of Charter Schools 2000* , National Study of Charter Schools, Fourth Year Report, (U.S. Department of Education, Office of Educational Research and Improvement, 2000). The survey asked principals to estimate students on free and reduced lunch, a broader category than free lunch alone. CC reports numbers for free and reduced lunch separately, but before 1998 combined them in one variable. I use the free lunch variable from 2000, which the NCES describes as comparable to the data collected before 1998.

[20] Simeon P. Slovacek, Antony J. Kunnan, Hae-Jin Kim, "California Charter Schools Serving Low-SES Students: An Analysis of the Academic Performance Index," (Charter College of Education, California State University, Los Angeles, 2002). Center for Education Reform, *National Charter School Directory, 2001-2002*, (Center for Education Reform, 2002).

[21] School of Urban and Public Affairs, University of Texas at Arlington, et al., *Texas Open-Enrollment Charter Schools Fourth-Year Evaluation*, (University of Texas at Arlington, 2001).

[22] Thomas J. Kane and Douglas O. Staiger, "Volatility in School Test Scores: Implications for Test-Based Accountability Systems," in *Brookings Papers on Education Policy 2002*, Diane Ravitch, Ed., (Brookings Institution Press, 2002), pp. 235-284.

[23] RPP International, *The State of Charter Schools 2000.*

[24] Data on whether schools offer a specialized curriculum were also gathered, but in no case did the variable reach significance. The analyses are not presented in the paper.

[25] Education Week, *Quality Counts 2001: A Better Balance; Standards, Tests, and The Tools to Succeed*, (January, 2001); RPP International, *The State of Charter Schools 2000.*

BEST COPY AVAILABLE

[26] I was not able to acquire school level responses from RPP International in time for this draft of the paper, only state means. Thus, the estimates in the "RPP" column in Table 4 were made by assigning state means to all charter schools in the regression.

[27] Gain scores were generated by first computing adjusted 1999 and 2001 scores, then subtracting the 1999 score from the 2001 score. Scores were adjusted by regressing test scores on race and SES variables and a charter dummy, then producing a predicted score for each school. Means were then computed, weighted for enrollment.

[28] Kane and Staiger (2002) found that gain scores almost appear random in NC and CA, as did Figlio and Page (2001) in FL. For a review of the concerns with value-added accountability systems, see Helen Ladd and Janet Hansen, Eds., *Making Money Matter: Financing America's Schools,* (National Academy Press, 1999).

[29] Charter schools must overcome several daunting challenges in the first few years, see Tom Loveless and Claudia Jasin, "Starting from Scratch: Organizational and Political Obstacles Facing Charter Schools," *Educational Administration Quarterly*, 34 (1), (February, 1998), pp. 9-30 .

[30] The demographic data for at-risk charters: African Americans, 28%; Hispanics, 55%; free lunch, 62%. For non at-risk charters: African Americans, 29%; Hispanics, 31%, free lunch, 45%.

[31] School of Urban and Public Affairs, et al., 2001.

[32] In 1999-2000 school year, the number of general open-enrollment charter schools was capped at 120, but state law allowed for an unlimited number of 75 Percent Rule schools (School of Urban and Public Affairs, et al., 2001).

[33] A Horace Mann conversion school may be granted a charter jointly by the local school committee, local teachers union, and the state board of education.

[34] Bruno Manno, Chester E. Finn, and Greg Vanourek, "Charter School Accountability: Problems and Prospects," *Educational Policy*, 14 (4), (September, 2000), pp. 473-493.

[35] Three of the closed MN schools were in low-income neighborhoods in St. Paul and Minneapolis and one served a 95% Native American population in Dakota.

[36] For an interesting newspaper account of regulations stifling charter schools, see Jonathan Goldstein and Tim Simmons, "Charters Go Back in the Box," *The News and Observer*, February 26, 2002. Available at www.newsobserver.com .

[37] See Bettinger (1999) for different findings from Hoxby's, that MI charters lag in achievement gains.

BEST COPY AVAILABLE

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

ERIC®

REPRODUCTION RELEASE

(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title:	Charter School Achievement and Accountability. PEPG/02-09	EA 032 098
Author(s):	Loveless, Tom	
Corporate Source:		Publication Date: 2002

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic media, and sold through the ERIC Document Reproduction Service (EDRS). Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following three options and sign at the bottom of the page.

The sample sticker shown below will be affixed to all Level 1 documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

1

Level 1

☑

Check here for Level 1 release, permitting reproduction and dissemination in microfiche or other ERIC archival media (e.g., electronic) *and* paper copy.

The sample sticker shown below will be affixed to all Level 2A documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE, AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

2A

Level 2A

☐

Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only

The sample sticker shown below will be affixed to all Level 2B documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

2B

Level 2B

☐

Check here for Level 2B release, permitting reproduction and dissemination in microfiche only

Documents will be processed as indicated provided reproduction quality permits.
If permission to reproduce is granted, but no box is checked, documents will be processed at Level 1.

I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche or electronic media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.

Sign here, → please

Signature: Tom S. Loveless	Printed Name/Position/Title: Tom S. Loveless	
Organization/Address: Brookings 1775 Massachusetts Ave NW, Washington, DC 20036	Telephone:	FAX:
	E-Mail Address: tloveless@brookings.edu	Date: 1/28/03

(over)

III. DOCUMENT AVAILABILITY INFORMATION (FROM NON-ERIC SOURCE):

If permission to reproduce is not granted to ERIC, *or*, if you wish ERIC to cite the availability of the document from another source, please provide the following information regarding the availability of the document. (ERIC will not announce a document unless it is publicly available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents that cannot be made available through EDRS.)

Publisher/Distributor:
Address:
Price:

IV. REFERRAL OF ERIC TO COPYRIGHT/REPRODUCTION RIGHTS HOLDER:

If the right to grant this reproduction release is held by someone other than the addressee, please provide the appropriate name and address:

Name:
Address:

V. WHERE TO SEND THIS FORM:

Send this form to the following ERIC Clearinghouse:

ERIC Clearinghouse on Educational Management
~~1787 Agate Street~~
5207 University of Oregon
Eugene, OR 97403-5207
Attn: Acquisitions

However, if solicited by the ERIC Facility, or if making an unsolicited contribution to ERIC, return this form (and the document being contributed) to:

ERIC Processing and Reference Facility
4483-A Forbes Boulevard
Lanham, Maryland 20706

Telephone: 301-552-4200
Toll Free: 800-799-3742
FAX: 301-552-4700
e-mail: info@ericfac.piccard.csc.com
WWW: http://ericfacility.org

EFF-088 (Rev. 2/2000)